

What Is Claimed:

1. A method of calculating a distance value between terms in a document comprising:

locating implicitly defined semantic structures in the document; and

calculating the distance values in the document based on the implicitly defined semantic structures.
2. The method of claim 1, wherein the document is a HTML (Hyper-Text Markup Language) document.
3. The method of claim 2, wherein the implicitly defined semantic structures include lists created with HTML tags.
4. The method of claim 3, wherein the HTML tags include paragraph tags, new line tags, bold tags, or table tags.
5. The method of claim 1, further comprising:

locating explicitly defined semantic structures.
6. The method of claim 1, wherein the semantic structures include lists.

7. The method of claim 1, wherein the distance values are calculated as a word count between pairs of terms in the document augmented by rules related to the implicitly defined semantic structures.

8. The method of claim 1, wherein locating the implicitly defined semantic structures includes:

locating repeating occurrences of a set of two or more text formatting commands.

9. The method of claim 1, wherein the implicitly defined semantic structures include titles or headings.

10. A device comprising:
means for locating implicitly defined semantic structures in a document;
and
means for calculating a distance value between terms in the document based on the implicitly defined semantic structures.

11. The device of claim 10, further comprising:
means for generating a ranking score for the document, based on the distance value, that defines a relevancy of the document to the terms.

12. A method for ranking documents relative to a search query, the method comprising:

determining semantically based distance values between terms that occur in the search query and that are present in the documents; and

ranking the documents for relevancy to the search query based on the determined distance values.

13. The method of claim 12, wherein determining the semantically based distance values includes:

determining whether one or more of the terms are present within a list.

14. The method of claim 13, wherein the list is implicitly defined.

15. The method of claim 13, wherein determining the semantically based distance values further includes:

assigning a distance value indicative of closeness when two terms are present in a same item of the list.

16. The method of claim 12, further comprising:
locating implicitly defined semantic structures in the documents; and
using the implicitly defined semantic structures in determining the semantically based distance values.

17. The method of claim 16, wherein the implicitly defined semantic structures are located prior to the ranking.

18. The method of claim 12, wherein the documents are HTML (Hyper-Text Markup Language) documents.

19. The method of claim 18, wherein the implicitly defined semantic structures include lists created with HTML tags.

20. The method of claim 19, wherein the HTML tags include paragraph tags, new line tags, bold tags, or table tags.

21. The method of claim 12, wherein determining the semantically based distance values includes:

determining whether one or more of the terms are present within a title or heading.

22. A device comprising:

a memory; and

a processor coupled to the memory to:

determine semantically based distance values between terms that occur in a search query and that are present in documents; and

rank the documents for relevancy to the search query based on the determined distance values.

23. The device of claim 22, wherein the processor further:

locates implicitly defined semantic structures in the documents; and

uses the implicitly defined semantic structures in determining the semantically based distance values.

24. The device of claim 22, wherein the processor further:
receives a search query that contains the terms.

25. A method comprising:
receiving a search query;
locating implicitly defined semantic structures in documents;
calculating distance values for the documents based on the implicitly defined semantic structures and based on terms in the search query;
ranking the documents for relevancy to the search query based on the distance values; and
presenting the documents in an order influenced by the ranking.

26. The method of claim 25, wherein the documents are HTML (Hyper-Text Markup Language) documents.

27. The method of claim 26, wherein the implicitly defined semantic structures include lists created with HTML tags.

28. The method of claim 25, further comprising:
locating explicitly defined semantic structures.